Segmentation of Scenes in Exploratory Mode
Final Report

PI Name:	Ruzena Bajcsy
Co-PI Name:	Max Mintz
PI Institution:	Department of Computer and Information Science, University of Pennsylvania
PI Phone Number:	(215) 898-0370
PI Fax Number:	(215) 573-2048
PI E-mail Address:	bajcsy@central.cis.upenn.edu

Grant Title:	Segmentation of Scenes in Exploratory Mode
Grant/Contract Number:	N00014-88-K-0630
R&T Number:	4331742-01
Reporting Period:	01 Oct 91 – 30 Sep 92

1 Objective

GRASP Laboratory research combines Active Perception and Robotics to produce intelligent devices capable of performing sophisticated tasks. This research specifically concentrates on multi-spectral image processing, 3-D shape identification, decision making and robot actuation. Perception via manipulation is combined with information obtained from a variety of sensors to establish one or more features or properties of an unstructured environment. This links exploration of an unknown environment by visual sensing, range measurement, manipulation and physical probing. It is a direct application of our theoretical work in robust multisensor fusion and techniques for integrating data from multiple modalities.

One of the primary objectives of this research is to investigate *coordination* and *communication* of multi-agent systems. In particular, multiple agents explore and adapt to their surroundings and organize and configure themselves to perform required tasks with possible assistance of human agents.

2 Approach

Our approach is based on an "advice-based small-team" architecture. The agents are heterogeneous in both their scope of applicability or functionality and their capabilities or competence. Their connectivity depends on their shared versus independent domain of applicability and/or task/subtasks. An example of a shared domain is an obstacle monitored by vision and acoustic sensors. These two agents perform redundant or complimentary functions. On the other hand, an example of independent agents would be the force sensors that monitor the contact and sliding/rolling of an object held by two palms, while the acoustic sensors monitor the vehicles to avoid obstacles. The advice-based small-team architecture is new in that it provides as much autonomy as possible to individual agents, yet it makes all the possible information accessible to other agents in the spirit of cooperation. All the agents know the common task of transporting an object from place A to place B. Since all the agents are physical, the real time issue becomes apparent!

3 Progress

Progress in the last year of the project has been made in the areas of control, the observer agent, and multisensor fusion.

3.1 Control

The control of individual mobile manipulators has been investigated. The control of a mobile manipulator involves the coordination of locomotion of the mobile platform and manipulation of the manipulator. The coordination of locomotion and manipulation is important for a number of reasons including redundancy in mobility, difference in dynamic response, nonholonomic constraints, and dynamic interactions. Modeling the mobile platform as a nonholonomic dynamic system, we have developed and experimentally tested a control algorithm for coordinating locomotion and manipulation. Using this algorithm, while the manipulator is dragged by an operator in any direction in a horizontal plane, the mobile platform is able to bring the manipulator into the configuration with maximum manipulability measure.

In further research we will integrate the wrist force/torque sensor in the control algorithm, thus enabling the mobile manipulator to maintain contact with and follow a moving surface rather than being dragged.

3.2 Observer Agent

The function of the Observer Agent is to recognize the environment (in real time), in particular the free path and the obstacles. While there are numerous algorithms in the literature describing how to extract optical flow, range and motion parameters, most of them are too complex to run in real time (15 frames per second) and not robust enough. For the real-time processing we have concentrated on proper data reduction mechanisms (data selection) via the use of different optics, and model-based tracking. For the robustness question, we have concentrated on removing the highlights and shadows using color and active light. These algorithms are based on point transformations (difference of two images), and hence are highly parallel.

3.3 Multisensor Fusion

Our multi-agent system employs the following sensors: multiple cameras which simultaneously provide images from multiple fields of view and varying depths of field; digital compasses on the mobile agents; odometry from the wheels of the mobile agents; acoustic range sensors on the mobile agents; infrared proximity sensors on the mobile agents; and force and torque sensors on the end-effectors of manipulator agents. These sensors provide information of different types and qualities. One research issue is delineation of decision models for combining or fusing information with a common type with differing qualities, as well as the fusion of information of different types with varying quality. We have developed a mathematical model for fusing information of a common type where one source provides coarse-grained information with good reliability and the other source provides fine-grained information but may be subject to serious sporadic errors. For example, we can use optics or infrared technology for coarse range determination. Within specified domains of operation, these coarse range estimates will be reliable; we can use acoustic range sensing for fine-grained range information — with the caveat that the acoustic range information may be seriously

in error due to multipath or insufficient target cross-section. These models and techniques do not rely on either highly refined sensor noise models or highly accurate sensor position information. Both of these additional sources of errors are accounted for in this methodology.

4 Accomplishments

Accomplishments in the last year of the project include:

- Recognition of highlights for dialectric materials and metallic materials
- Recognition of shadows using active light.
- Simultaneous real-time (15 frames per second) model-based 2D tracking of multiple objects.
- Near-optimal robust fixed-size confidence procedures.
- Robustness with respect to noise distribution uncertainty, applicable to essentially any class of noise distributions which have smooth (non-atomic) boundaries.
- Near-optimal performance obtained using easily computed non-monotone functions.

5 Technical Reports

1. Tarek M. Sobh and Ruzena Bajcsy. *Visual Observation Of A Moving Agent.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-91-86, GRASP LAB 283.
2. Marcos Salganicoff and Ruzena Bajcsy. *Sensorimotor Learning Using Active Perception In Continuous Domains.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-91-87, GRASP LAB 284.
3. Eric Paljug, Tom Sugar, Vijay Kumar and Xiaoping Yun. *Important Considerations In Force Control With Applications To Multi-Arm Manipulation.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-91-88, GRASP LAB 287.
4. Sanjay Agrawal. *Robotic Manipulation Using A Behavioral Framework.* Technical Report (Dissertation), Department of Computer and Information Science, University of Pennsylvania, MS-CIS-91-90, GRASP LAB 287.
5. Ruzena Bajcsy and Mario Campos. *Active and Exploratory Perception.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-91-91, GRASP LAB 288.
6. Ruzena Bajcsy. *An Active Observer.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-91-95, GRASP LAB 295.
7. Tarek Sobh. *Active Observer: A Discrete Event Dynamic System Model For Controlling An Observer Under Uncertainty.* Technical Report (Dissertation), Department of Computer and Information Science, University of Pennsylvania, MS-CIS-91-99, GRASP LAB 296.

8. Gareth D. Funka-Lea. *Vision For Navigation Using Two Road Cues.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-91-100, GRASP LAB 297.

9. Thomas Lindsay. *Teleprogramming: Remote Site Research Issues.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-01, GRASP LAB 298.

10. John Bradley. *Interactive Image Display For The X Window System.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-04, GRASP LAB 299.

11. Luca Bogoni. *Superquadric Library, User Manual and Utility Programs.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-11, GRASP LAB 300.

12. Pramath Raj Sinha. *Robotic Exploration Of Surfaces and Its Application To Legged Locomotion.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-12, GRASP LAB 301.

13. Sang Wook Lee. *Understanding Of Surface Reflections In Computer Vision By Color and Multiple Views.* Technical Report (Dissertation), Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-13, GRASP LAB 301.

14. Faculty & Graduate Students. *Grasp Laboratory News, Volume 8, Number 1.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-15, GRASP LAB 302.

15. Yin-Tien Wang and Vijay Kumar. *Simulation Of Mechanical Systems With Multiple Frictional Contacts.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-16, GRASP LAB 303.

16. Yoshio Yamamoto and Xiaoping Yun. *Coordinating Locomotion and Manipulation Of A Mobile Manipulator.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-18, GRASP LAB 304.

17. Eric D. Paljug. *Multi-Arm Manipulation Of Large Objects With Rolling Contacts.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-19, GRASP LAB 305.

18. Insup Lee. *Proving Properties of Real-Time Distributed Systems: A Comparison of Three Approaches.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-20, GRASP LAB 306.

19. Robert Bruce King, II. *Design, Implementation, and Evaluation Of A Real-Time Kernel For Distributed Robotics.* Technical Report (Dissertation), Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-26, GRASP LAB 307.

20. Marcos Salganicoff. *A Robotic System for Learning Visually-Driven Grasp Planning.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-27, GRASP LAB 308.

21. Gerda L. Kamberova. *Robust Location Estimation for MLR and Non-MLR Distributions.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-28, GRASP LAB 309.

22. Gerda L. Kamberova. *Markov Random Field Models: A Bayesian Approach To Computer Vision Problems.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-29, GRASP LAB 310.

23. Robert Mandelbaum. *Convergence of Stochastic Processes.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-30, GRASP LAB 311.

24. Gerda Kamberova, Ray McKendall and Max Mintz. *Multivariate Data Fusion Based On Fixed-Geometry Confidence Sets.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-31, GRASP LAB 312.

25. Jana Košecká. *Control of Discrete Event Systems.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-35, GRASP LAB 313.

26. Luca Bogoni and Ruzena Bajcsy. *An Active Approach To Functionality Characterization and Recognition.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-37, GRASP LAB 315.

27. Mario Fernando Montenegro Campos. *Robotic Exploration Of Material and Kinematic Properties Of Objects.* Technical Report (Dissertation), Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-38, GRASP LAB 316.

28. Ron Katriel. *Parallel Evidence-Based Indexing of Complex Three-Dimensional Models Using Prototypical Parts and Relations.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, MS-CIS-92-39, GRASP LAB 317.

6 Publications

1. Ruzena Bajcsy. An Active Observer. *Proceedings of the ARPA Image Understanding Workshop,* pages 137–147, San Diego, CA, January 1992.

2. Mario Campos, Vijay Kumar and Ruzena Bajcsy. Kinematic identification of linkages. *Proceedings of the 3rd International Conference on Advances in Robot Kinematics*, Ferrara, Italy, September 7-9, 1992.

3. Janez Funda, Thomas Lindsay and Richard P. Paul. Teleprogramming: towards delay-invariant remote manipulation. *Presence: Teleoperators and Virtual Environments*, Volume 1, Number 1; MIT Press, January 1992.

4. Gareth Funka-Lea and Ruzena Bajcsy. Vision for vehicle guidance using two road cues. *Proceedings of the Intelligent Vehicles '92 Symposium*, pages 126–131, Detroit, Michigan, June 1992.

5. Alok Gupta, Gareth Funka-Lea and Kwangyoen Wohn. Segmentation, modeling and classification of the compact objects in a pile. Hatem Nasr, editor, *Selected Papers on Automatic Object Recognition*, SPIE Milestone Series, 1991.

6. Gerda Kamberova, Ray McKendall and Max Mintz. Multivariate Data Fusion Based on Fixed-Geometry Confidence Sets. *SPIE Proc. of the International Symposium on Advances in Intelligent Systems, Session on Sensor Fusion*, November 1991.

7. V. Koivunen and M. Pietikainen. Evaluating quality of surface description using robust methods. *11th International Conference on Pattern Recognition*, The Hague, Netherlands, pp. 214-218, 1992.

8. V. Koivunen and M. Pietikainen. Experiments with combined edge and region-based range image segmentation. *Theory and Applications of Image Analysis*, World Scientific Publications, pp. 162-176, 1992.

9. Vijay Kumar. Characterization of workspaces of parallel manipulators. *ASME Journal of Mechanical Design*, 114(3):368–375, 1992.

10. Vijay Kumar. A compact inverse velocity solution for redundant robots. In *Proceedings of 1992 International Conference on Robotics and Automation*, pages 482–487, Nice, France, May 1992.

11. V. Kumar, T. G. Sugar and G. Pfreundschuh. A three degree of freedom in-parallel actuated manipulator. *Proceedings of the 9th CISM-IFToMM Symposium on Theory and Practice of Manipulators*, Udine, Italy, September 1-4, 1992.

12. Sang Wook Lee and Ruzena Bajcsy. Detection of specularity using color and multiple views. *Image and Vision Computing*, 10:643–653, 1992.

13. Sang Wook Lee and Ruzena Bajcsy. Detection of Specularity Using Color and Multiple Views. *Proc. of Second European Conference on Computer Vision*, Santa Margherita Ligure, Italy, 1992. Outstanding Paper Award.

14. Ales Leonardis and Ruzena Bajcsy. Finding parametric curves in an image. *Proc. of Second European Conference on Computer Vision*, Santa Margherita Ligure, Italy, 1992.

15. Jasna Maver and Ruzena Bajcsy. Occlusions and the Next View Planning. *IEEE Int. Conf. on Robotics and Automation*, May 1992.

16. Ray McKendall and Max Mintz. Robust Sensor Fusion with Statistical Decision Theory. *Data Fusion in Robotics and Machine Intelligence*, M.A. Abidi and R.C. Gonzalez, editors, Academic Press, Spring, 1992.

17. Mohamed Ouerfelli, William Harwin and Vijay Kumar. A pneumatic actuation system for a wheelchair-mounted robot arm. *15th RESNA International Conference*, Toronto, June 6-11, 1992.

18. Eric Paljug, Thomas Sugar, Vijay Kumar and Xiaoping Yun. Important considerations in force control with applications to multi-arm manipulation. *IEEE International Conference on Robotics and Automation*, pp. 1270-1275, Nice, France, May 10-15, 1992.

19. Richard Paul, Thomas Lindsay and Craig Sayers. Time delay insensitive teleoperation. *Proceedings of the 1992 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 247–254, July 1992.

20. Richard Paul, Thomas Lindsay, Craig Sayers and Matt Stein. Time-delay insensitive virtual-force reflecting teleoperation. *Artificial Intelligence, Robotics and Automation in Space*, pages 55–67, Toulouse, France, September 1992.

21. Marcos Salganicoff and Ruzena Bajcsy. Robot sensorimotor learning in continuous domains. *Proceedings of 1992 International Conference on Robotics and Automation*, Nice, France, May 1992.

22. Pramath R. Sinha and Ruzena K. Bajcsy. Implementation of an Active Perceptual Scheme for Legged Locomotion of Robots. *Proceedings of the Fourth International Workshop on Intelligent Robots and Systems (IROS '91)*, pages 1518–1523, Osaka, Japan, November 1991.

23. Pramath R. Sinha and Ruzena K. Bajcsy. Robotic Exploration of Surfaces and its Application to Legged Locomotion. *Proceedings of the IEEE International Conference on Robotics and Automation*, Nice, France, May 1992.

24. Tarek M. Sobh and Ruzena Bajcsy. A Model for Observing a Moving Agent. *Proceedings of the Fourth International Workshop on Intelligent Robots and Systems (IROS '91)*, Osaka, Japan, November 1991.

25. Tarek M. Sobh and Ruzena Bajcsy. A Model for Visual Observation Under Uncertainty. *1992 IEEE Symposium on Computer Aided Control System Design (CACSD '92)*, March 1992.

26. Tarek M. Sobh and Ruzena Bajcsy. Autonomous Observation Under Uncertainty. *IEEE International Conference on Robotics and Automation*, Nice, France, May 1992.

27. Matt Stein and Richard Paul. Kinesthetic replay for error diagnosis in time delayed teleoperation. *SPIE OE/Technology '92: Telemanipulator Technology*, 1992.

28. Chau-Chang Wang, Nilanjan Sarkar and Vijay Kumar. Rate kinematics of mobile manipulators. *Proceedings of the 22nd Biennial ASME Mechanisms Conference*, pages 225–232, Scottsdale, AZ, September 1992.

29. Y. Wang and V. Kumar. Simulation of mechanical systems with unilateral constraints. *Proceedings of the 22nd Biennial ASME Mechanisms Conference*, Scottsdale, AZ, September 1992.

30. Y. Wang, V. Kumar, and J. Abel. Dynamics of rigid bodies undergoing multiple frictional contacts. *IEEE International Conference on Robotics and Automation*, pp. 2764-2769, Nice, France, May 10-15 1992.

31. Yangsheng Xu, Xiaoping Yun and Richard P. Paul. Nonlinear feedback control of robot manipulator and compliant wrist. *Dynamics and Control*, (1):325–339, 1991.

32. Xiaoping Yun. Modeling and control of two constrained manipulators. *Journal of Intelligent and Robotic Systems*, (4):363–377, 1991.

33. Xiaoping Yun. Nonlinear feedback for force control of manipulators. C.T. Leondes, editor, *Control and Dynamic Systems*, pages 259–283, Academic Press, New York, 1991.

34. Xiaoping Yun and Daizhan Cheng. Input-output decoupled linearization of general nonlinear systems. *Transactions of the Institute of Measurement and Control*, 13(4):218–224, 1991.

35. Xiaoping Yun and Vijay Kumar. An approach to simultaneous control of trajectory and interaction forces in dual arm configurations. *IEEE Transactions on Robotics and Automation*, 7(5):618–625, October 1991.

36. Xiaoping Yun, Vijay Kumar, Nilanjan Sarkar and Eric Paljug. Control of multiple arms with rolling constraints. *1992 IEEE International Conference on Robotics and Automation*, pages 2193–2198, Nice, France, May 1992.

An Active Observer

Ruzena Bajcsy*
GRASP Laboratory
Computer and Information Science Department
University of Pennsylvania
Philadelphia, PA 19104

1 Abstract

In this paper we present a framework for research into the development of an **Active Observer**. The components of such an observer are the low and intermediate visual processing modules. Some of these modules have been adapted from the community and some have been investigated in the GRASP laboratory, most notably modules for the understanding of surface reflections via color and multiple views and for the segmentation of three dimensional images into first or second order surfaces via superquadric/parametric volumetric models. However the key problem in Active Observer research is the control structure of its behavior based on the task and the situation. This control structure is modeled by a formalism called Discrete Events Dynamic Systems (DEDS).

2 Introduction

We are interested in the development of an Active Observer. An Active Observer is an agent which has capabilities to observe scenes, objects, situations and deliver the observed information to human, manipulatory, and mobile agents. Naturally there are more questions than answers. We shall list a few which are of particular interest to us. What are the components/modules that such an observer must have? How are these components interconnected, i.e. what is the architecture of such an agent? Some of the modules correspond to certain visual cues. We take as a given that our observer has several such cues. In that case, the subsequent question is how are the results from these cues integrated? When are they invoked? How is the selection process conducted/guided? Which cue is employed and when? Finally, what kind of information/messages is delivered by the observer to other agents?

Towards this end, for the last two years we have concentrated on the development of theoretical and experimental understanding of some of the cues/components, some cues' integration and selection, and control strategies for observation capability. In particular, in cue development we have tried to understand surface reflections by color and multiple views. An important finding of this work, which will be described in detail in Section 2, is that multiple view points provide useful information for discriminating between specular and Lambertian reflections both from dielectrics and from metals. In Section 3, we shall describe a system for the segmentation of a three dimensional scene into components that can be modeled by superquadric parametric fit. This system uses, in cooperation, surface segmentation, contour segmentation and gross volumetric segmentation in order to arrive at the proper result. The scenes are of moderate complexity (up to 10 parts), but no other assumptions are made about objects or their parts. This work points to the common fact that one module or cue or approach cannot handle the perceptual variety of the data that the real world, even in moderate complexity, represents. Multiple cues are necessary and hence a great deal of thought has to go into the integration policy and control structure. In Section 4, we present a formal model of an observer agent. This model is based on the theory of Discrete Event Dynamic Systems (DEDS), which allows us to unequivocally predict the observation capabilities of an observer. In order for this to occur, the observer must know the discrete events of the task. So far this is done by the designer. Finally, in Section 5 we show the recent development of a CCD chip (the Retina) with space variant resolution. Details are described in this section.

3 Understanding of Reflection Properties Using Color and Multiple Views

Recently there has been a growing interest in the detection of specularity in both basic and applied computer vision research. In general, the detection of specularities from a single gray-level image is a physically underconstrained problem, and more information needs to be collected in physically sensible ways to solve the problem. Successful development of an algorithm for image data collection and interpretation necessarily depends on physical models that describe how surfaces appear according to the illumination and reflectance properties and sensor characteristics. Recently the computer vision field has increasingly incorporated methodologies derived from physical principles of image formation and

*Acknowledgement: Navy Grant N0014-88-K-0630, AFOSR Grants 88-0244, AFOSR 88-0296; Army/DAAL 03-89-C-0031PRI; NSF Grants CISE/CDA 88-22719, IRI 89-06770; and Dupont Corporation

sensing [7]. So far there have been three types of approaches to solving the problem of specularity detection through the collection of more images: (1) with different light directions, (2) with different sensor polarization angles, and (3) with different color sensors.

The photometric-stereo-type approaches consider the specular and Lambertian reflectance properties for obtaining object shape using more than two light directions [4] [9] [11]. Since the direction and the degree of the collimation of the illumination need to be strictly controlled, application of the approach is restricted to dark-room environments. The polarization method analyzes the polarization of reflected light and detects specularities from dielectrics and metals [12]. The polarization approach places some restrictions on the incident illumination direction with respect to surface orientation.

The dichromatic model [10] proposed by Shafer has been the key model to the recent specularity detection algorithms using color [8] [5] [6] [3]. The basic limitation of the color algorithms is that objects must be only colored dielectrics to use the dichromatic model. For color image segmentation, it is usually assumed that object surface reflectance is spatially piecewise uniform in color and that scene illumination is singly colored. We have previously developed a color image segmentation algorithm for the separation of diffuse, as well as sharp, specularities and inter-reflections from Lambertian reflections [3].

Our recent research has focused on the development of some specularity detection or separation methods that only require modification of sensors but not any modification of environments. In other words, they are methods that are active in modifying sensors but passive in modifying environments. There are two kinds of modification of environments: relocation and re-orientation of objects by robot manipulation, and illumination change. The prime example of the illumination change is the light switching for the photometric-stereo-type methods. Since illumination lighting needs to be strictly controlled, the photometric-stereo-type approaches are applicable only for inspection in dark rooms.

Strict illumination control is not always possible in investigating surface reflection properties in many general environments. Examples include outdoor inspection, indoor or outdoor navigation, and exploratory environments. Even for indoor inspection, a well controlled dark room is not always available.

For general environments without strict illumination control, only sensors are controllable, and color and polarization can be the possible cues. Another possibility is to move the observer, which has not been used for investigating reflection properties in computer vision. The idea of moving the observer was directly motivated by the concept of active vision [2]. For low-level vision problems of shape or structure, it has been demonstrated that many ill-posed problems become well-posed if more information is collected by active sensors [1]. Although the paradigms for shape or structure based on feature correspondence cannot be directly applied to the study of reflectance properties, the idea of a moving observer motivated the investigation of new principles by physical modeling in obtaining more information.

In this paper, we suggest the use of multiple views for the detection of specularity by introducing two algorithms. The first algorithm, called spectral differencing, uses color information from a small number of multiple views. The second algorithm is called view sampling. Using many views of gray-level images collected in wide angle, the view sampling reconstructs object structure and detects specularities. An important principle used for the algorithms is the Lambertian consistency, which is the well-known fact that the Lambertian reflection does not change its brightness and spectral content depending on viewing directions, but the specular reflection or the mixture of Lambertian and specular reflections can change.

A problem associated with the use of multiple views with color is what kind of extra spectral information can be obtained by moving a color camera without considering object geometry. If there is any, it may alleviate the limiting assumptions imposed on the object and illumination domain for the color segmentation approaches, and provide higher confidence in detecting specularities.

The spectral differencing algorithm is based on the observation that any presence of specular reflections can be inferred by the difference in the distribution of pixel colors between two color images. According to the Lambertian consistency, the color distribution of pixels from only Lambertian reflections should be consistent regardless of view points. On the other hand, specularities or the mixture of specular and Lambertian reflections can change the distribution of pixel colors between two views.

The spectral differencing algorithm does not require any assistance from image segmentation and geometrical manipulation. Since the algorithm does not rely on the segmentation and the dichromatic model, it is applicable to dielectric objects with nonuniform reflectance and metals under multiply colored illumination. Figures 1 and 2 show two dielectric objects with variation in reflectance and a metallic object in neutral reflectance color. Two fluorescent light tubes and a tungsten light bulb are used for illumination and there are inter-reflections from the walls. $MSD(0 \leftarrow 1)$ shows the regions of new color distribution in view 0 compared to view 1, and $MSD(1 \leftarrow 0)$ the regions of new color distribution in view 1 compared to view 0. Under multiply colored and extended illumination, it can be seen that most of the specularities are detected by the spectral differencing.

Another approach we introduce is to obtain reflection properties using only multiple views without any color information. With densely sampled views in wide angle and with known viewing directions, the view sampling algorithm reconstructs object structure as well as detects specularities from Lambertian reflections. The view sampling algorithm is applicable to dielectrics and metals.

If object structure is reconstructed assuming the Lambertian consistency for both Lambertian and specular reflections, the structure reconstructed from the specular reflections would not in general represent the real object

Figure 1: Spectral differencing

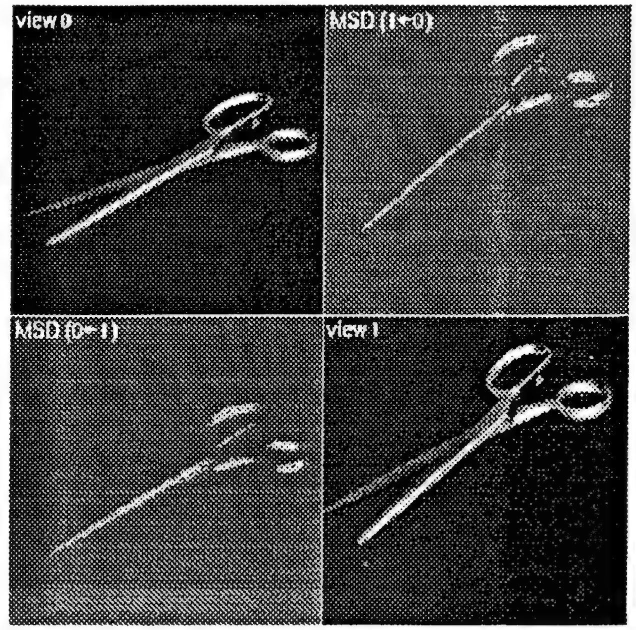

Figure 2: Spectral differencing

surface, while the one reconstructed from the Lambertian reflections does. By examining the differently reconstructed object structures from specular and Lambertian reflections, we can identify the reflection types and the real object structure.

We adopted an algorithm for computerized tomography through photometric modeling for the reconstruction of object structure. Figure 3 shows the camera control scheme and Figure 4 (a) shows 4 out of 30 view samples of a gray dielectric object from different view points. Figure 4 (c) and (d) show the reconstructed structures at the cross sections 1 and 2 illustrated in Figure 4 (b), respectively. As shown in Figure 4 (c) and (d), the structure reconstructed from specularities at the cross section 2 is different from the real object surface reconstructed by Lambertian reflections.

The future direction of our studies is the integration of many cues in the light of active vision [2]. Active vision involves not only the modeling of physical sensing and data processing for vision modules (local model), but also the control of the modules (global model). Global models characterize the overall performance and make predictions on how the individual modules will interact, which in turn determines how intermediate results are combined. It is the global model that analyzes and combines the information from many visual cues to assign stable descriptors. For more stable descriptions of reflection properties in more general environments, it is desirable to extract extra information from a synergistic combination of multiple cues. The spectral differencing algorithm demonstrates the synergy from the combination of color and multiple views. There are also potentials for extra information from the combination of color, polarization and multiple views.

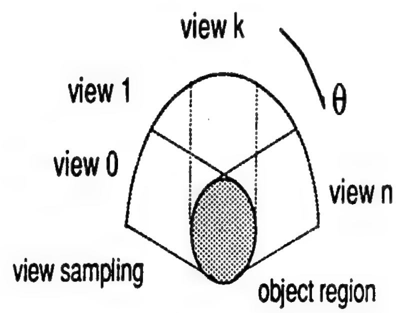

Figure 3: View sampling

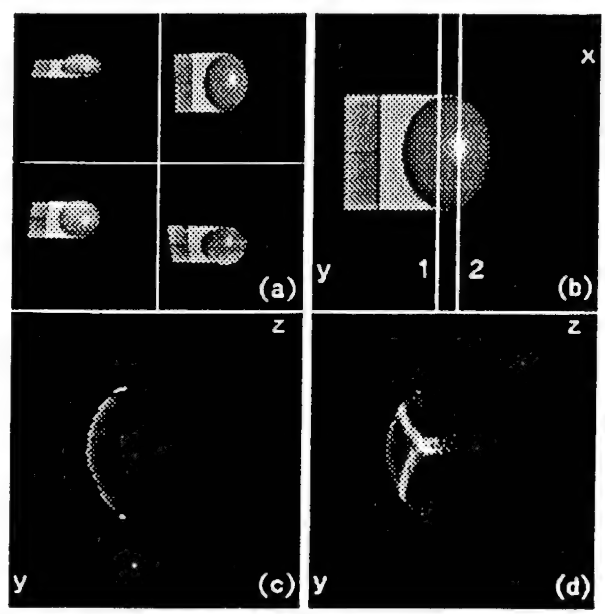

Figure 4: View sampling

4 Surface and Volumetric Segmentation of Complex 3-D Objects Using Parametric Shape Models

The problem of part definition, description, and decomposition is central to shape recognition systems. In this paper we present an integrated framework for segmenting dense range data of complex 3-D scenes into their constituent parts in terms of surface (bi-quadrics) and volumetric (superquadrics) primitives, without a priori domain knowledge or stored models. Our objective is to *recover* a structured description of complex 3-D objects, guided entirely by the geometric properties of the shape models. The resulting decomposition into parts is very useful for the high-level processes, which can attach domain specific labels to the parts, and reason at a level where the visual input is structured in terms of geometric primitives, rather than cope with the difficulties of low-level vision and a huge amount of unstructured data.

Since the shapes have to be recovered from raw data, it is not possible to invoke complex models (models with hundreds of degrees of freedom) straight away. It is, however, feasible and perceptually less ambiguous to use simpler but powerful models that can capture the local and global properties of the object shapes, and provide a first approximation to the more complex models. With computability, simplicity, and the utility of the shape representation as our major concerns, we use bi-quadrics and superquadrics as our surface and volumetric models respectively. We develop SUPERSEG (SUPERquadric SEGmentation), a control structure to effectively carry out the decomposition of complex objects in range images, and address the numerous issues encountered in a data-driven bottom-up approach [13; 14; 15].

The SUPERSEG system 5 has five major components: namely, the bi-quadric surface segmentation module; the module for extracting surface properties and adjacency relationships; the superquadric model recovery module; the residual generation and analysis module; and the control module for superquadric-based segmentation.

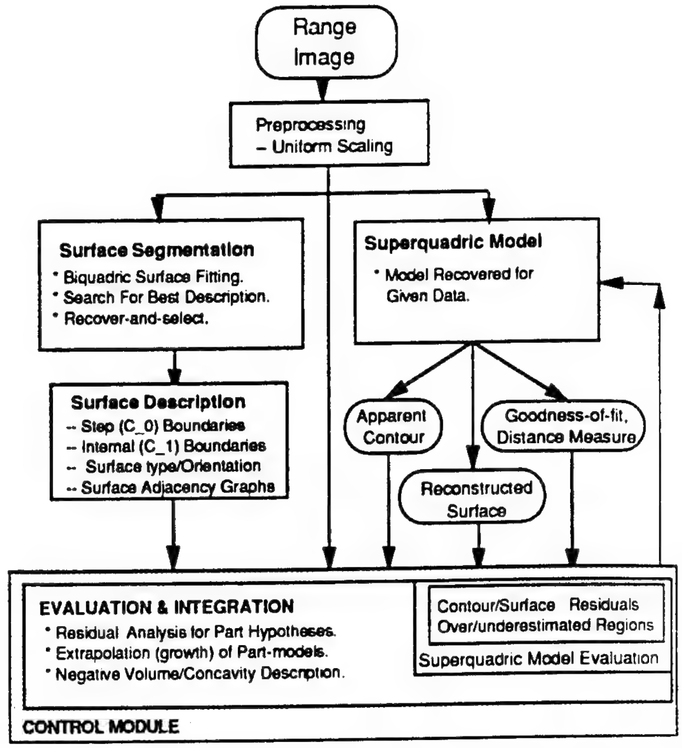

Figure 5: The SUPERSEG system: A framework for surface and volumetric segmentation.

4.1 Surface Segmentation: Bi-quadric Models

The surface segmentation is performed by a novel local-to-global iterative regression approach of searching for the best piecewise description of the data in terms of bi-quadric models [16; 17]. The model-recovery module consists of independently extrapolating all the seed-regions and fitting the model using the least-squares regression method. The region-growing is controlled by a *compatibility-constraint*, whose value depends on the noise due to sensor and quantization, as well as the allowed tolerance between the shapes of the model and underlying data. Seed-regions are placed in a grid-pattern all over the image, and allowed to grow until they are either completely grown or rejected by the model-selection procedure (which maximizes a linear benefit-cost function). Instead of first growing all the regions and then invoking the model-selection procedure (Recover-then-select), the model-recovery and model-selection processes are dynamically combined (Recover-and-select) to achieve a computationally feasible and robust method capable of rejecting outliers and determining its domain of applicability.

4.1.1 Refining Surface Segmentation & Extracting Surface properties

The bi-quadric segmentation achieved by the above procedure needs refinement before it can be used as an intermediate segmentation by superquadric-based volume segmentation. Also, the coefficients of the second-order surfaces have information about orientation and surface-type (convex or concave) inherent in them. The orientation information is tremendously useful in aligning the major axis of cylindrical superquadric models. Further, due to the compatibility-constraint, regions intersecting to form surface normal discontinuities (C_1) overlap in the vicinity of the discontinuity, thereby localizing it. We developed a systematic method for tracing the biquadric intersection curve, which is used to refine the segmentation as well as to localize the discontinuities (edges) and to characterize them as convex or concave. In addition, a surface adjacency graph (SAG) is constructed with surface patches as nodes and discontinuity-type as edges between them. The information extracted from the bi-quadric patches is used to generate and test hypotheses by the volumetric segmentation module.

4.2 Superquadrics: Volumetric Part-Models

Superquadric models are convex part-models (except the bent models) that can be recovered for a given set of 3-D points by minimizing a function based on the modified implicit inside-outside superquadric function [18; 19;

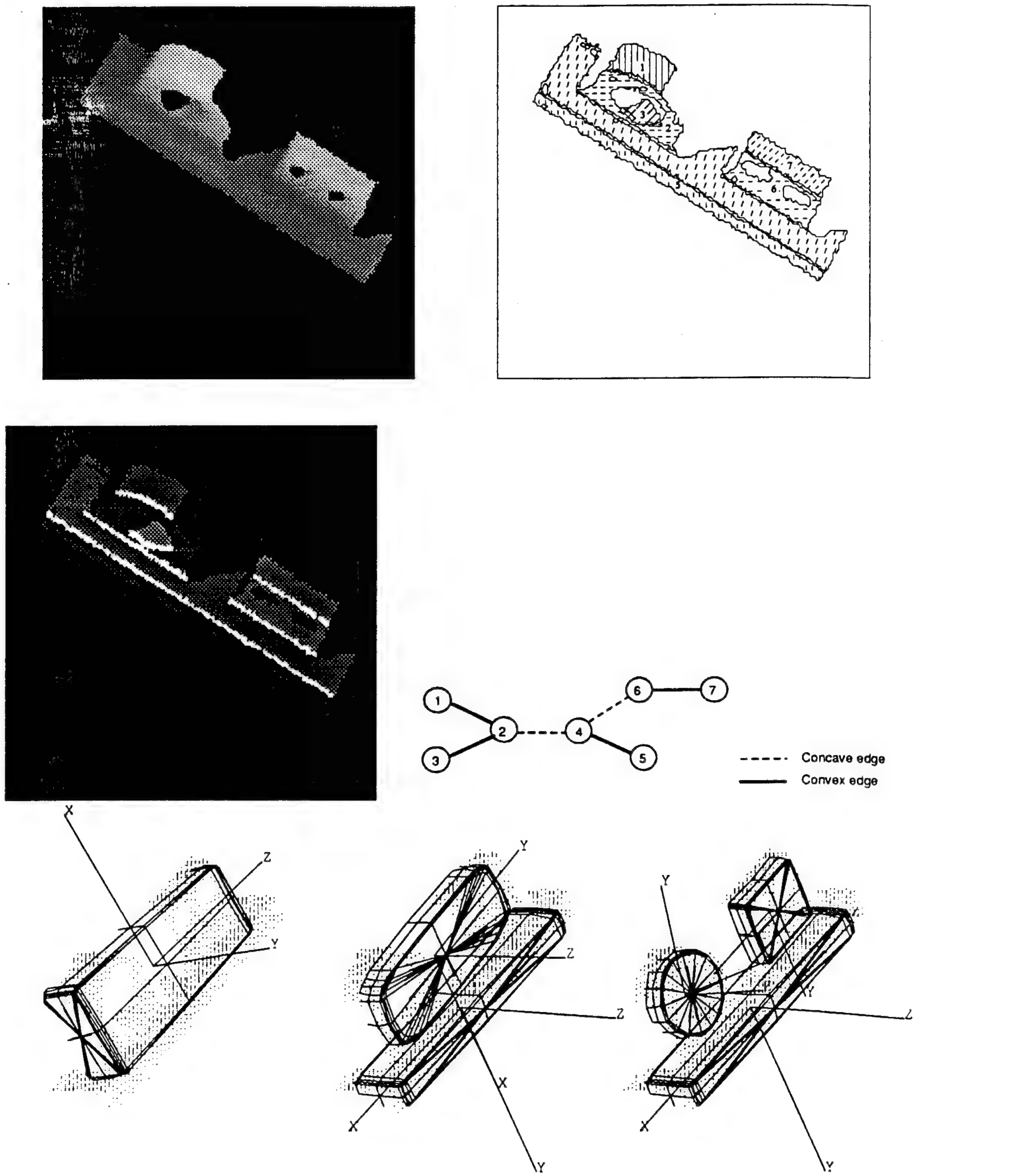

Figure 6: **The NIST object:** Top: The range image and its bi-quadric surface segmentation. Center: the C_1 (surface normal) edges marked at the overlapping parts of the surfaces. Following a procedure similar to the intersection cleaning, the edges are marked as convex or concave and a surface adjacency graph (SAG) is constructed. Bottom: The three iterations of the global-to-local procedure to extract the part-structure.

15]. This formulation enforces a minimum volume constraint as well as a surface constraint, but is incapable of decomposing the data set if no appropriate convex model can be found in the model vocabulary. Thus, the superquadric model recovery module is adequate only for recovering an optimal model (if oriented correctly) given a data set, but not for segmenting it. To decide whether a recovered model is adequate for the given data set, we have developed an exhaustive set of criteria comprised of qualitative and quantitative measures. Quantitative measures are the normalized global deviation of the model from data. The deviation can be the inside-outside function value, or can be measured along the direction of the viewpoint (Z-residuals for a range scanner), or along the direction of the minimum distance of a point from the model (Euclidean distance). The qualitative measures are the 'local' residuals characterized by the clusters of 3-D points that are either inside the model, or on the model, or outside the model. Both qualitative and quantitative measures are necessary for complete evaluation of a recovered model.

4.3 Volumetric Segmentation: The Control Strategy

In view of the fact that volumetric models don't have good surface support (as opposed to bi-quadric models), they cannot be recovered by following exclusively the extrapolation method (local-to-global) used by bi-quadrics. In order to obtain an optimal piecewise-convex volumetric segmentation, it is necessary to proceed global-to-local, where data is decomposed only if the global model is inadequate. This allows controlled residual-driven decomposition of 3-D data, as also introduction of an objective evaluation criteria for an acceptable description. However, the global-to-local method can be aided by the bi-quadric segmentation in forming hypotheses about convex combination of surfaces, which although is not true in general (an L shape for example), can significantly reduce the computational overhead if true for a particular part. Previous researchers have assumed that a 1-to-1 mapping exists between surface patches and superquadric models, which is also not true in general. But it does provide a planarity check for the patches, as well as the orientation and shape of the individual patches in 3-space.

Thus, a strategy that combines the bi-quadric information with the global-to-local residual-driven method is most effective in recursively segmenting the scene to derive the part-structure [13]. A set of acceptance criteria based on the quantitative and qualitative measures provide the objective evaluation of intermediate descriptions, and decide whether to terminate the procedure, or selectively refine the segmentation, or generate negative volume description. The control module generates hypotheses about superquadric models at clusters of underestimated data and performs controlled extrapolation of part-models by shrinking the global model. The recursive splitting of data results in a hierarchical part-structure comprising of global and local models. The results of complete processing of the range image of a machined object (from NIST) is shown in Figure 6.

We have tested the SUPERSEG system on real range images of scenes of varying complexity, including objects with occluding parts, and scenes where surface segmentation is not sufficient to guide the volumetric segmentation. Some of the applications of our approach include data reduction, 3-D object recognition, geometric modeling, automatic model generation, object manipulation, qualitative vision, and active vision.

5 A Framework for Visual Observation

In this work we establish a framework for the general problem of observation, which may be applied to different kinds of visual tasks. We define "intelligent" high-level control mechanisms for the observer in order to achieve efficiency in recognizing different processes within a specific dynamic system. The intelligent observer is able to recognize the visual tasks, understands the meaning of the scene evolution and successfully reports on the current visual state. It is obvious that there is a need for high-level interpretation of actions within the environment and to have guarantees for observation capabilities and stability within the viewing mechanism. The framework is a *predictable* one that satisfies the following general requirements:

- Recognizes visual tasks and events.
- Repositions itself adaptively and intelligently.
- Operates in real time.
- Asserts and reports on *distinct* and *discrete* visual states.
- Utilizes the *continuous* parametric evolution of the visual system.
- Accommodates visual uncertainties.

We concentrate on observing a manipulation process in order to illustrate the ideas and motive behind our framework. The process of observing a robot hand manipulating an object is very crucial for many robotic and manufacturing tasks. It is important to know in an automated manufacturing environment whether the robot hand is doing the correct sequence of operations on an object (or more than one object). It might be a fact that the workspace of the robotic manipulator cannot be accessed by humans, as in the case of some space applications or some areas within a nuclear plant, for example. In such a case, having another robot "look" at the process is a very good option. Thus, the observation process can be thought of as a stage in a closed-loop fully automated system where there are robots who perform the required manipulation task and some other robots who observe them and correct their actions when something goes wrong. Typical manipulation processes include grasping, pushing, pulling, lifting, squeezing, screwing and unscrewing. In this work, we address the problem of observing a single hand manipulating a single object and recognizing what the hand is doing. No feedback will be supplied to the manipulating robot to correct its actions. We divide the problem into three major components. First, we identify a high-level framework for the visual states. Next, we define the events that cause

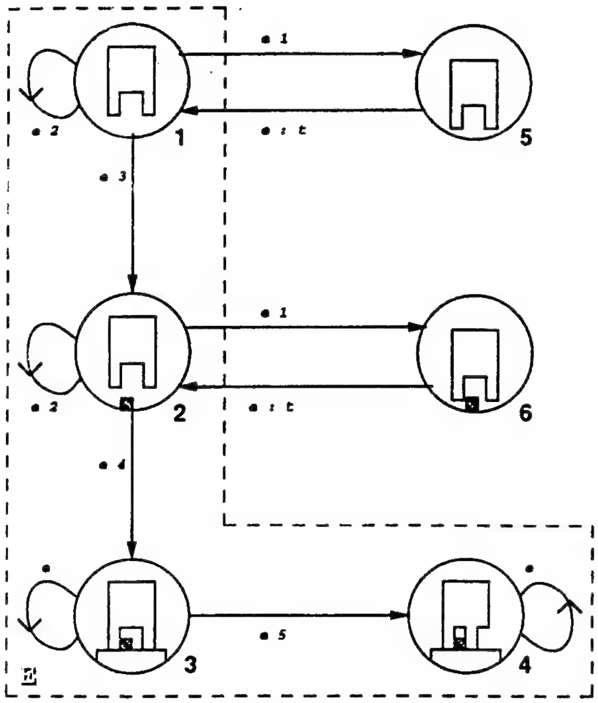

Figure 7: A Model for a Grasping Task

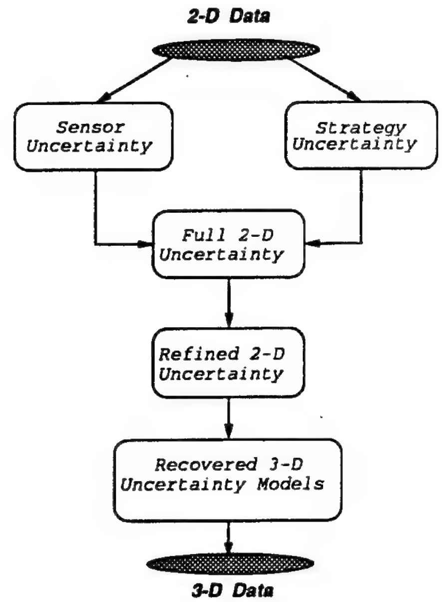

Figure 8: Propagation of Uncertainty

state transitions. Finally, we utilize visual uncertainties to assert the state of the system.

5.1 State Space Modeling

We use a discrete event dynamic system as a high-level structuring skeleton to model the visual manipulation system. Discrete event dynamic systems (DEDS) are dynamic systems (typically asynchronous) in which state transitions are triggered by the occurrence of discrete events in the system. Our formulation uses the knowledge about the system and the different actions in order to solve the observer problem in an efficient, stable and practical way. The model incorporates different hand/object relationships and the possible errors in the manipulation actions. It also uses different tracking mechanisms so that the observer can keep track of the workspace of the manipulating robot. A framework is developed for the hand/object interaction over time and a stabilizing observer is constructed. The construction process utilizes a task-dependent coarse quantization of the manipulation actions in order to attain an active, adaptive and goal-directed sensing mechanism. An example of a DEDS automaton for a simple grasping task is shown in Figure 7.

5.2 Event Identification

Low-level modules are developed for recognizing the "events" that cause state transitions within the dynamic manipulation system. To be able to observe how the events evolve over time, we must be able to identify how the hand moves and how the hand/object physical relationship evolves over time. We use a mix of 2-D and 3-D modules to recover a set of parameters that define the *continuous* parametric evolution of the scene under observation. Three dimensional evolution of the hand motion is recovered by tracking a set of features and two-dimensional cues to the number of objects and their relative location; two dimensional motion with respect to the manipulating hand is recovered in real-time. The recovered events are then used to assert state transitions within the DEDS automata. We also recover uncertainties associated with the visual event recovery and utilize them for navigating the observer automata.

5.3 Utilizing Uncertainties

This work examines closely the possibilities for errors, mistakes and uncertainties in the visual manipulation system, observer construction process and event identification mechanisms. We divide the problem into a number of major *levels* for developing uncertainty models in the observation process. The propagation of uncertainty is shown in Figure 8.

The *sensor* level models deal with the problems in mapping 3-D features to pixel coordinates and the errors incurred in that process. We identify these uncertainties and suggest a framework for modeling them. The next level is the *extraction strategy* level, in which we develop models for the possibility of errors in the low-level image processing modules used for identifying features that are to be used in computing the 2-D evolution of the scene under consideration. In the following level, we utilize the geometric and mechanical properties of the hand and/or object to reject unrealistic estimates for 2-D movements that might have been obtained from the first two levels. We transform the 2-D uncertainty models into 3-D uncertainty models for the structure and motion of the entire scene. The next level uses the equations that

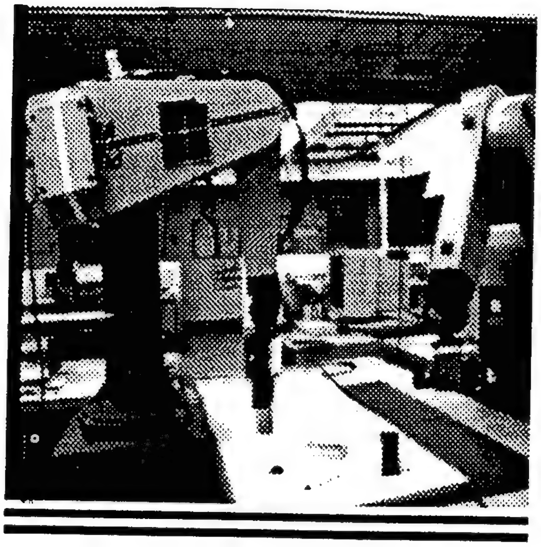

Figure 9: Experimental Setting

govern the 2-D to 3-D relationship to perform the conversion. We then reject the improbable 3-D uncertainty models for motion and structure estimates by using the existing information about the geometric and mechanical properties of the moving components in the scene. The highest level is the DEDS formulation with uncertainties, in which state transitions and event identification is asserted according to the 3-D models of uncertainty that were developed in the previous levels, and error recovery is performed according to the ordering of the recovered distributions.

5.4 Conclusions

The approach used can be considered as a framework for a variety of visual tasks, as it lends itself to be a practical and feasible solution that uses existing information in a robust and modular fashion. The work examines closely the possibilities for errors and uncertainties in the manipulation system, observer construction process and event identification mechanisms. Ambiguities are allowed to develop and are resolved after finite time; recovery mechanisms are devised too. Details of the observer system can be found in [20; 21; 22; 23]. Theoretical and experimental aspects of the work support adopting the framework as a new basis for performing task-oriented recognition, inspection and observation of visual phenomena. The observer and manipulating robots experimental setup is shown in Figure 9.

6 Spatio-Variant Sensing

Traditional imaging for robotics vision has relied almost exclusively on common commercial imagers, notably television format sensors. Their advantages are clear: the cameras are inexpensive and readily available, and the sampling of the data is on a "natural" cartesian (x,y) grid. These sensors have placed enormous demands, however, on processing architectures. The problem is not only that image analysis is an ill-defined task in the real world, but that we have only very expensive machines that can begin to process the data.

Over the last seven years an international team, led by Van der Spiegel at the University of Pennsylvania, Sandini at DIST in Italy, and Claeys at IMEC in Belgium, designed, built, and tested a new imaging chip called the Retina [24]. The new camera serves as the foundation to a new approach to robotics vision. We shift the focus at the systems level from gathering better data and designing machines to analyze it to gathering data for the computing resources that exist. The result is a prototype sensor that reduces the computational complexity of the problem by three orders of magnitude and, if scaled to commercial cameras, by six orders [25].

The Retina attempts to model the gross characteristics of the primate visual system in a mathematically elegant way. The computational savings arise from the same mechanism the eye uses, namely, to maintain one area of high resolution on the focal plane and to drop the resolution elsewhere. The mathematical expression of this is a log-polar mapping. That mapping transforms a polar data space, where a point P has the polar coordinates $(r, theta)$, by taking the logarithm of the expression for the point:

$$P = re^{i\theta} \longrightarrow P'^{ln} = ln(r) + i\theta = u + iv$$

This mapping has the useful property of separating rotations (changes in *theta*) from magnifications (changes in r). If the sensor has a uniform sampling grid in u (ln (r)), then the spatial grid in r will exponentially grow as distance from the center grows. This models the growth of the receptive fields in primate retinas.

The Retina layout in Figure 10 implements this mapping by sampling in $(r, theta)$ at points matching a uniform (u, v) grid. The sensor clearly has rotational symmetry and exponentially decreasing resolution. The circular section contains only 1920 pixels (30 circles of 64 pixels/circle); at the center is a dense rectangular grid of 102 additional photosites [26]. The cells grow fast: the outermost circle is over ten times as wide as the innermost. This leads directly to the small pixel count.

The chip, with its custom driving electronics, is now working at the GRASP laboratory [27] and is producing good pictures as shown in Figure 11.

Clearly visible in the data space is the large magnification of the inner circles. The outer section provides much poorer data, with pixels widely spaced and averaging the incident light over a larger area. Still they do not provide useless information.

The nature of the information has changed, however. No longer do we get high quality data across the focal plane. Indeed, we assume from the start that we do not try to build a model of the world in one step. Instead, we use the periphery to guide our attention—where we point the camera. Implicit here is the idea of an active observer. The Retina, just sitting on a bench waiting for an object to enter its high-resolution spot, is useless. We must actively build the world by moving the camera, using the periphery to suggest candidates for attention.

The cost of using this sensor might be considered high. The new data space will require rewriting or adapting

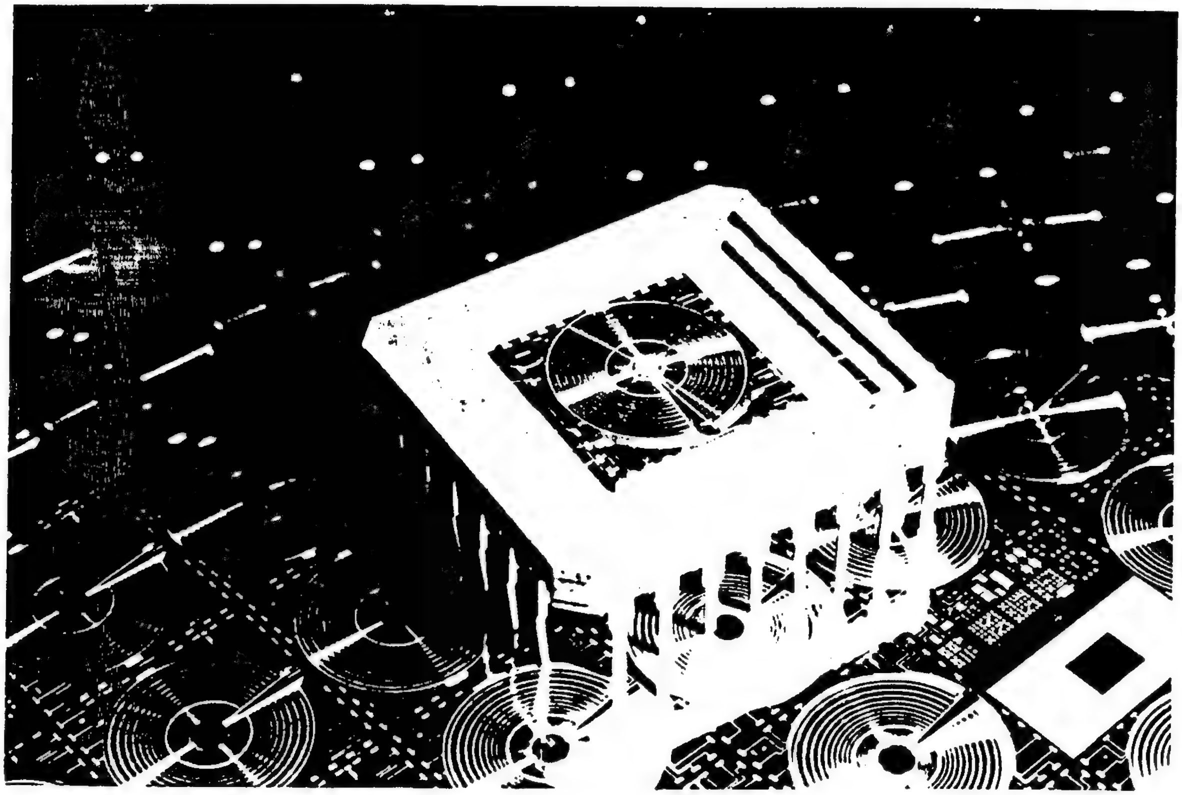

Figure 10: The Retina CCD Imager

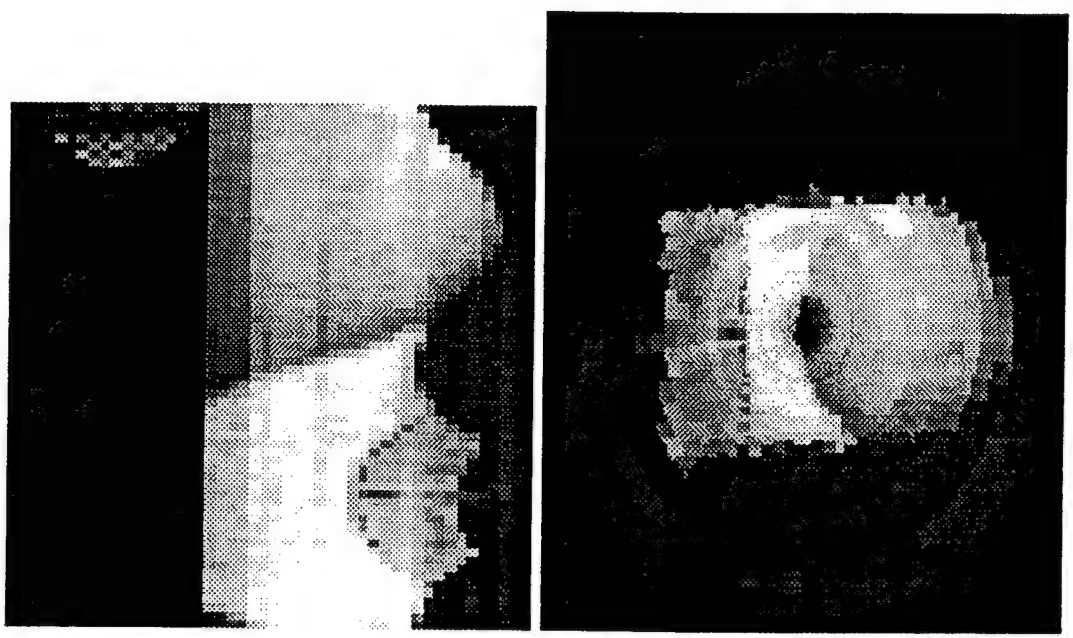

Figure 11: Picture of a mouse from the camera, centered between the buttons (to the left) and ball. The picture on the left is in the mapped plane: the vertical axis is v (v, the angle of the point, increases moving down the axis) and the horizontal is u (u, the log of the radial distance of the point, increases to the right). The triangle at the upper left of the image is data remapped back onto a cartesian grid.

all our tools for the cartesian plane: this is the primary cost outside the hardware development. The advantages, however, suggest profit. The Retina has some one hundred times fewer pixels than a standard television camera, which drastically reduces the computational burden of analysis, bringing it within the abilities of modern machines. The gains also include the rich mathematical structure of the mapping. That structure simplifies pattern matching by making rotations and magnifications linear shifts in the data space, and speeds time-to-impact measurements by looking only at a radial flow. Some distortions introduced by the mapping, such as translational variance (which is linear translations becoming curves in the data space) also disappear in an active observer, where for example attention and tracking automatically compensate for linear motion.

Since the sensor began working this summer, our focus at the GRASP laboratory has been redeveloping traditional image processing tools. Our work has looked at edge detection in the new data space, detecting lines using a Hough algorithm, calculating the centroid of an object, and measuring time-to-impact. Each of these areas requires an analysis of their mathematical basis under the log mapping and coding the results on real images. All algorithms must further be computationally simple to work in a real-time environment.

This integration of sensor and computer is now the fundamental area of research involving the Retina at Penn. That the Retina works proves the concept of the hardware, of designing custom imaging sensors for robots. The integration itself will prove the concept of the system. The Retina is the basic building block for a real-time interactive observer.

7 Conclusions and future plans

The development of an Active Observer is underway at the GRASP laboratory. Although future emphasis will be placed on the control structure of such an observer, its integration policies, and communication issues with other observers and agents in general, there is still a need for further studies, developments and improvements of component technologies. For example, in the case of understanding surface reflectance, we still have not completed the theoretical underpinning of transparency. With the problem of segmentation, while the cooperation between surface and volumetric fittings is necessary, and they help in resolving ambiguities, the first and second order primitives are clearly not sufficient for modeling a broad class of real life objects. Higher order models will have to be invoked, but only selectively and locally after the lower order fits have failed. If this order of fitting data is violated then instabilities in the fitting procedures can be expected. Finally, there is the question of the control mechanism of the Active Observer. As shown above, we have employed the Discrete Event Dynamic System model. DEDS is a suitable formalism to model continuous processes of observation, as well as events occuring in discrete intervals. As a result, this model allows us to predict the observation capability as defined by the control theory community. The assumption here, however, is that the task of observation is *a priori* in terms of the discrete events. While in the original theory the transitions from one state/event to another were discrete, we have extended the theory to transitions with uncertainties. The next task should be to loosen the requirements for explicit knowledge of the desired observable events. These events should be able to be generated from some rules of physics, geometry and other conventions of the object's and agent's interactions. In conclusion, we are on our way to complete an Active Observer which has a control structure that allows us to predict observation capabilities. The components developed here allow the Active Observer to handle moderately complex scenes of shapes/materials, their spatial arrangements and their illuminations. The real time issue of processing is a crucial one and hence our efforts in special purpose CCD chips and related hardware. The open questions are many but we wish to concentrate on the intercommunication of several observers and other agents, such as manipulatory, mobile and human agents. Ultimately, the final issue is this: who tells what and how much, and to whom.

References

[1] J. Aloimonos and A. Badyopadhyay. Active vision. In *Proc. 1st Int. Conf. on Computer Vision*, pages 35–54, 1987.

[2] R. Bajcsy. Active perception. *Proceedings of the IEEE*, 76(8):996–1005, 1988.

[3] R. Bajcsy, S.W. Lee, and A. Leonardis. Color image segmentation with detection of highlights and local illumination induced by inter-reflections. In *Proc. 10th International Conf. on Pattern Recognition*, Atlantic City, NJ, June 1990.

[4] E. N. Coleman and R. Jain. Obtaining 3-dimensional shape of textured and specular surface using four-source photometry. *Computer Graphics and Image Processing*, 18(4):308–328, 1982.

[5] R. Gershon. *The Use of Color in Computational Vision*. PhD thesis, Department of Computer Science, University of Toronto, 1987.

[6] G.H. Healey and T.O. Binford. Using color for geometry-insensitive segmentation. *Journal of the Optical Society of America*, 6, 1989.

[7] T. Kanade and K. Ikeuchi. Introduction to the special issue on physical modeling in computer vision. *IEEE Trans. PAMI*, 13(7):609–610, 1991.

[8] G.J. Klinker, S.A. Shafer, and T. Kanade. Image segmentation and reflection analysis through color. In *Proceedings of the DARPA Image Understanding Workshop*, pages 838–853, Pittsburgh, PA, 1988.

[9] S. K. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Trans. Robo. Autom.*, 6(4):418–431, 1990.

[10] S.A. Shafer. Using color to separate reflection components. *COLOR Research and Application*, 10(4):210–218, 1985.

[11] H.D. Tagare and R. J. deFigueiredo. Photometric stereo for diffuse non-lambertian surface. *IEEE Trans. PAMI*, 13():, 1991.

[12] L. B. Wolff. *Polarization Methods in Computer Vision*. PhD thesis, Department of Computer Science, Columbia University, 1991.

[13] Gupta Alok, *Surface and Volumetric Segmentation of Complex 3-D Objects Using Parametric Shape Models*. Technical Report MS-CIS-91-45, Department of Computer and Information Science, University of Pennsylvania, 1991.

[14] Gupta, Alok and Ruzena Bajcsy, Part description and segmentation using contour, surface and volumetric primitives, in *Proceedings of the Conference on Sensing and Reconstruction of 3D Objects and Scenes*, pp. 203–214, SPIE, Santa Clara, CA, Feb 1990.

[15] Gupta, Alok, Luca Bogoni, and Ruzena Bajcsy, Quantitative and qualitative measures for the evaluation of the superquadric models, in *Proceedings of the IEEE Workshop on Interpretation of 3D Scenes*, pp. 162–169, Austin, TX, November 1989.

[16] Leonardis, Ales, Alok Gupta, and Ruzena Bajcsy, Segmentation as the search for the best description of the image in terms of primitives, in *Proceedings of the Third International Conference on Computer Vision*, pp. 121–125, IEEE, Osaka, Japan, December 1990a.

[17] Leonardis, Ales, Alok Gupta, and Ruzena Bajcsy, *Segmentation of Range Images as the Search for the Best Description of the Scene in Terms of Geometric Primitives*, Technical Report MS-CIS-90-30, CIS Department, University of Pennsylvania, 1990b.

[18] Solina, Franc, *Shape Recovery and Segmentation with Deformable Part Models*, PhD thesis, University of Pennsylvania, 1987, Technical Report MS-CIS-87-111.

[19] Solina, F. and R. Bajcsy, Recovery of parametric models from range images: the case for superquadrics with global deformations, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, *12*, 131–147, February 1990.

[20] R. Bajcsy and T. M. Sobh, *A Framework for Observing a Manipulation Process*. Technical Report MS-CIS-90-34 and GRASP Lab. TR 216, Computer Science Dept., School of Engineering and Applied Science, University of Pennsylvania, June 1990.

[21] T. M. Sobh, *A Framework for Visual Observation*. Technical Report MS-CIS-91-36 and GRASP Lab. TR 261, Computer Science Dept., School of Engineering and Applied Science, University of Pennsylvania, May 1991.

[22] T. M. Sobh and R. Bajcsy, "Visual Observation of A Moving Agent". In *Proceedings of the European Robotics and Intelligent Systems Conference (EURISCON '91)*, Corfu, Greece, June 1991 and presented at the 12^{th} *International Joint Conference on Artificial Intelligence (IJCAI)*, Workshop on Dynamic Scene Understanding, Sydney, Australia, August 1991.

[23] T. M. Sobh and R. Bajcsy, "A Model for Observing a Moving Agent". In *Proceedings of the Fourth International Workshop on Intelligent Robots and Systems (IROS '91)*, Osaka, Japan, November 1991.

[24] J. Van der Spiegel, G. Kreider, et al. "A Foveated Retina-Like Sensor Using CCD Technology". In *Analog VLSI Implementations of Neural Systems*, ed. C. Mead and M. Ismail, pp. 189-211, Kluwer Academic Publishers, Boston, 1989.

[25] G. Kreider, J. Van der Spiegel et al. "The Design and Characterization of a Space Variant CCD Sensor". SPIE Vol. 1381 *Intelligent Robots and Computer Vision IX: Algorithms and Techniques*, Boston, November 1990.

[26] G. Kreider, J. Van der Spiegel et al. "A Retina-Like Space Variant CCd Sensor". SPIE Vol. 1242 *Charge Coupled Devices and Solid State Optical Sensors*, pp. 133-140, Santa Clara, February 1990.

[27] Z. Kalayjian. "A Driver Circuit for the Foveated Retina-Like Optical Sensor". Final Report, Undergraduate Fellowship in Sensor Technologies, University of Pennsylvania, 1990.

OFFICE OF THE UNDER SECRETARY OF DEFENSE (ACQUISITION)
DEFENSE TECHNICAL INFORMATION CENTER
CAMERON STATION
ALEXANDRIA, VIRGINIA 22304-6145

IN REPLY REFER TO DTIC-OCC

SUBJECT: Distribution Statements on Technical Documents

TO: OFFICE OF NAVAL RESEARCH
CORPORATE PROGRAMS DIVISION
ONR 353
800 NORTH QUINCY STREET
ARLINGTON, VA 22217-5660

1995 1027 004

1. Reference: DoD Directive 5230.24, Distribution Statements on Technical Documents, 18 Mar 87.

2. The Defense Technical Information Center received the enclosed report (referenced below) which is not marked in accordance with the above reference.

FINAL REPORT
N00014-88-K-0630
TITLE: SEGMENTATION OF SCENES
IN EXPLORATORY MODE

3. We request the appropriate distribution statement be assigned and the report returned to DTIC within 5 working days.

4. Approved distribution statements are listed on the reverse of this letter. If you have any questions regarding these statements, call DTIC's Cataloging Branch, (703) 274-6837.

FOR THE ADMINISTRATOR:

1 Encl

GOPALAKRISHNAN NAIR
Chief, Cataloging Branch

FL-171
Jul 93

DISTRIBUTION STATEMENT A:

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION IS UNLIMITED

DISTRIBUTION STATEMENT B:

DISTRIBUTION AUTHORIZED TO U.S. GOVERNMENT AGENCIES ONLY; (Indicate Reason and Date Below). OTHER REQUESTS FOR THIS DOCUMENT SHALL BE REFERRED TO (Indicate Controlling DoD Office Below).

DISTRIBUTION STATEMENT C:

DISTRIBUTION AUTHORIZED TO U.S. GOVERNMENT AGENCIES AND THEIR CONTRACTORS; (Indicate Reason and Date Below). OTHER REQUESTS FOR THIS DOCUMENT SHALL BE REFERRED TO (Indicate Controlling DoD Office Below).

DISTRIBUTION STATEMENT D:

DISTRIBUTION AUTHORIZED TO DOD AND U.S. DOD CONTRACTORS ONLY; (Indicate Reason and Date Below). OTHER REQUESTS SHALL BE REFERRED TO (Indicate Controlling DoD Office Below).

DISTRIBUTION STATEMENT E:

DISTRIBUTION AUTHORIZED TO DOD COMPONENTS ONLY; (Indicate Reason and Date Below). OTHER REQUESTS SHALL BE REFERRED TO (Indicate Controlling DoD Office Below).

DISTRIBUTION STATEMENT F:

FURTHER DISSEMINATION ONLY AS DIRECTED BY (Indicate Controlling DoD Office and Date Below) or HIGHER DOD AUTHORITY.

DISTRIBUTION STATEMENT X:

DISTRIBUTION AUTHORIZED TO U.S. GOVERNMENT AGENCIES AND PRIVATE INDIVIDUALS OR ENTERPRISES ELIGIBLE TO OBTAIN EXPORT-CONTROLLED TECHNICAL DATA IN ACCORDANCE WITH DOD DIRECTIVE 5230.25, WITHHOLDING OF UNCLASSIFIED TECHNICAL DATA FROM PUBLIC DISCLOSURE, 6 Nov 1984 (Indicate date of determination). CONTROLLING DOD OFFICE IS (Indicate Controlling DoD Office).

The cited documents has been reviewed by competent authority and the following distribution statement is hereby authorized.

A

(Statement)

OFFICE OF NAVAL RESEARCH
CORPORATE PROGRAMS DIVISION
ONR 353
800 NORTH QUINCY STREET
ARLINGTON, VA 22217-5660

(Controlling DoD Office Name)

(Reason)

(Controlling DoD Office Address, City, State, Zip)

DEBRA T. HUGHES
DEPUTY DIRECTOR
CORPORATE PROGRAMS OFFICE

Debra T. Hughes

(Signature & Typed Name)

(Assigning Office)

19 SEP 1995

(Date Statement Assigned)